

1642

RECEIVED

APR 17 2001

TECH CENTER 1600/2900

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/509,775

DATE: 04/05/2001
 TIME: 11:09:33

Input Set : A:\53466277.app
 Output Set: N:\CRF3\04042001\I509775.raw

3 <110> APPLICANT: FUJITA, JUN
 5 <120> TITLE OF INVENTION: GANKYRIN
 7 <130> FILE REFERENCE: 053466/0277
 9 <140> CURRENT APPLICATION NUMBER: 09/509,775
 10 <141> CURRENT FILING DATE: 2000-03-31
 12 <150> PRIOR APPLICATION NUMBER: PCT/JP98/04467
 13 <151> PRIOR FILING DATE: 1998-10-02
 15 <150> PRIOR APPLICATION NUMBER: JP 9-286214
 16 <151> PRIOR FILING DATE: 1997-03-10
 18 <160> NUMBER OF SEQ ID NOS: 16
 20 <170> SOFTWARE: PatentIn Ver. 2.1
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 780
 24 <212> TYPE: DNA
 25 <213> ORGANISM: Homo sapiens
 27 <220> FEATURE:
 28 <221> NAME/KEY: CDS
 29 <222> LOCATION: (97)..(774)
 31 <400> SEQUENCE: 1

32 tgggtgaagct ctaacggctg ttttgactgg cgtagccgga gccggcgacg tgaggcgggc 60
 34 gttgctcgcg cgacaagtag ttgctgggac agcgaa atg gag ggg tgt gtg tct 114
 35 Met Glu Gly Cys Val Ser
 36 1 5
 38 aac cta atg gtc tgc aac ctg gcc tac agc ggg aag ctg gaa gag ttg 162
 39 Asn Leu Met Val Cys Asn Leu Ala Tyr Ser Gly Lys Leu Glu Glu Leu
 40 10 15 20
 42 aag gag agt att ctg gcc gat aaa tcc ctg gct act aga act gac cag 210
 43 Lys Glu Ser Ile Leu Ala Asp Lys Ser Leu Ala Thr Arg Thr Asp Gln
 44 25 30 35
 46 gac agc aga act gca ttg cac tgg gca tgc tca gct gga cat aca gaa 258
 47 Asp Ser Arg Thr Ala Leu His Trp Ala Cys Ser Ala Gly His Thr Glu
 48 40 45 50
 50 att gtt gaa ttt ttg ttg caa ctt gga gtg cca gtg aat gat aaa gac 306
 51 Ile Val Glu Phe Leu Leu Gln Leu Gly Val Pro Val Asn Asp Lys Asp
 52 55 60 65 70
 54 gat gca ggt tgg tct cct ctt cat att gcg gct tct gct ggc cgg gat 354
 55 Asp Ala Gly Trp Ser Pro Leu His Ile Ala Ala Ser Ala Gly Arg Asp
 56 75 80 85
 58 qaq att gta aaa gcc ctt ctg qga aaa ggt gct caa qtg aat gct gtc 402
 59 Glu Ile Val Lys Ala Leu Leu Gly Lys Gly Ala Gln Val Asn Ala Val
 60 90 95 100
 62 aat caa aat ggc tgt act ccc tta cat tat gca gct tcg aaa aac agg 450
 63 Asn Gln Asn Gly Cys Thr Pro Leu His Tyr Ala Ala Ser Lys Asn Arg
 64 105 110 115
 66 cat gag atc gct gtc atg tta ctg gaa ggc ggg gct aat cca gat gct 498
 67 His Glu Ile Ala Val Met Leu Leu Glu Gly Gly Ala Asn Pro Asp Ala
 68 120 125 130

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/509,775

DATE: 04/05/2001

TIME: 11:09:33

Input Set : A:\53466277.app

Output Set: N:\CRF3\04042001\I509775.raw

```

70 aag gac cat tat gag gct aca gca atg cac cgg gca gca gcc aag ggt 546
71 Lys Asp His Tyr Glu Ala Thr Ala Met His Arg Ala Ala Lys Gly
72 135 140 145 150
74 aac ttg aag atg att cat atc ctt ctg tac tac aaa gca tcc aca aac 594
75 Asn Leu Lys Met Ile His Ile Leu Leu Tyr Tyr Lys Ala Ser Thr Asn
76 155 160 165
78 atc caa gac act gag ggt aac act cct cta cac tta gcc tgt gat gag 642
79 Ile Gln Asp Thr Glu Gly Asn Thr Pro Leu His Leu Ala Cys Asp Glu
80 170 175 180
82 gag aga gtg gaa gaa gca aaa ctg ctg gtg tcc caa gga gca agt att 690
83 Glu Arg Val Glu Glu Ala Lys Leu Leu Val Ser Gln Gly Ala Ser Ile
84 185 190 195
86 tac att gag aat aaa gaa gaa aag aca ccc ctg caa gtg gcc aaa ggt 738
87 Tyr Ile Glu Asn Lys Glu Glu Lys Thr Pro Leu Gln Val Ala Lys Gly
88 200 205 210
90 ggc ctg ggt tta ata ctc aag aga atg gtg gaa ggt taaaca 780
91 Gly Leu Gly Leu Ile Leu Lys Arg Met Val Glu Gly
92 215 220 225
95 <210> SEQ ID NO: 2
96 <211> LENGTH: 226
97 <212> TYPE: PRT
98 <213> ORGANISM: Homo sapiens
100 <400> SEQUENCE: 2
101 Met Glu Gly Cys Val Ser Asn Leu Met Val Cys Asn Leu Ala Tyr Ser
102 1 5 10 15
104 Gly Lys Leu Glu Glu Leu Lys Glu Ser Ile Leu Ala Asp Lys Ser Leu
105 20 25 30
107 Ala Thr Arg Thr Asp Gln Asp Ser Arg Thr Ala Leu His Trp Ala Cys
108 35 40 45
110 Ser Ala Gly His Thr Glu Ile Val Glu Phe Leu Leu Gln Leu Gly Val
111 50 55 60
113 Pro Val Asn Asp Lys Asp Asp Ala Gly Trp Ser Pro Leu His Ile Ala
114 65 70 75 80
116 Ala Ser Ala Gly Arg Asp Glu Ile Val Lys Ala Leu Leu Gly Lys Gly
117 85 90 95
119 Ala Gln Val Asn Ala Val Asn Gln Asn Gly Cys Thr Pro Leu His Tyr
120 100 105 110
122 Ala Ala Ser Lys Asn Arg His Glu Ile Ala Val Met Leu Leu Glu Gly
123 115 120 125
125 Gly Ala Asn Pro Asp Ala Lys Asp His Tyr Glu Ala Thr Ala Met His
126 130 135 140
128 Arg Ala Ala Ala Lys Gly Asn Leu Lys Met Ile His Ile Leu Leu Tyr
129 145 150 155 160
131 Tyr Lys Ala Ser Thr Asn Ile Gln Asp Thr Glu Gly Asn Thr Pro Leu
132 165 170 175
134 His Leu Ala Cys Asp Glu Glu Arg Val Glu Glu Ala Lys Leu Leu Val
135 180 185 190
137 Ser Gln Gly Ala Ser Ile Tyr Ile Glu Asn Lys Glu Glu Lys Thr Pro
138 195 200 205

```

RAW SEQUENCE LISTING

DATE: 04/05/2001

PATENT APPLICATION: US/09/509,775

TIME: 11:09:33

Input Set : A:\53466277.app

Output Set: N:\CRF3\04042001\I509775.raw

```

140 Leu Gln Val Ala Lys Gly Gly Leu Gly Leu Ile Leu Lys Arg Met Val
141      210      215      220
143 Glu Gly
144 225
147 <210> SEQ ID NO: 3
148 <211> LENGTH: 696
149 <212> TYPE: DNA
150 <213> ORGANISM: Mus sp.
152 <220> FEATURE:
153 <221> NAME/KEY: CDS
154 <222> LOCATION: (1)..(693)
156 <220> FEATURE:
157 <221> NAME/KEY: modified_base
158 <222> LOCATION: (198)
159 <223> OTHER INFORMATION: a, t, c or g
161 <400> SEQUENCE: 3
162 atg gag ggg tgt gtg tct aac ata atg atc tgt aac ctg gcc tac agt 48
163 Met Glu Gly Cys Val Ser Asn Ile Met Ile Cys Asn Leu Ala Tyr Ser
164 1 5 10 15
166 ggg aag ctg gat gag ttg aag gag cgc att ttg gct gat aaa tct ctg 96
167 Gly Lys Leu Asp Glu Leu Lys Glu Arg Ile Leu Ala Asp Lys Ser Leu
168 20 25 30
170 gct act aga act gat cag qac agc aqa aca gct ttg cac tgg qca tgc 144
171 Ala Thr Arg Thr Asp Gln Asp Ser Arg Thr Ala Leu His Trp Ala Cys
172 35 40 45
174 tca gct ggc cat aca gaa att gtt gaa ttc ttg ctg caa ctt gga gtg 192
175 Ser Ala Gly His Thr Glu Ile Val Glu Phe Leu Leu Gln Leu Gly Val
176 50 55 60
W--> 178 cca gtn aat gat aaa gat gac gca ggt tgg tct cct ctt cat att gct 240
179 Pro Val Asn Asp Lys Asp Ala Gly Trp Ser Pro Leu His Ile Ala
180 65 70 75 80
182 gcc tcc gct ggc cgg gat gag att gta aaa gcc ctt ctg gtg aaa ggt 288
183 Ala Ser Ala Gly Arg Asp Glu Ile Val Lys Ala Leu Leu Val Lys Gly
184 85 90 95
186 gca cat gtt aat tct qtc aat caa aac gcc tgc act cca ctc cat tat 336
187 Ala His Val Asn Ser Val Asn Gln Asn Gly Cys Thr Pro Leu His Tyr
188 100 105 110
190 gca gct tcy aag aat agg cat gag att tct gtt atg tta cta gaa ggt 384
191 Ala Ala Ser Lys Asn Arg His Glu Ile Ser Val Met Leu Leu Glu Gly
192 115 120 125
194 ggg gct aac cca gat ggg aag gac cat tac gat gct aca gca atg cac 432
195 Gly Ala Asn Pro Asp Ala Lys Asp His Tyr Asp Ala Thr Ala Met His
196 130 135 140
198 egg gca gca gcc aag ggt aac ttg aag atg gtt cac atc ctt ctg ttc 480
199 Arg Ala Ala Ala Lys Gly Asn Leu Lys Met Val His Ile Leu Leu Phe
200 145 150 155 160
202 tac aaa gca tcc aca aac atc caa gac act gag ggt aac act cct cta 528
203 Tyr Lys Ala Ser Thr Asn Ile Gln Asp Thr Glu Gly Asn Thr Pro Leu
204 165 170 175

```

RAW SEQUENCE LISTING

DATE: 04/05/2001

PATENT APPLICATION: US/09/509,775

TIME: 11:09:33

Input Set : A:\53466277.app

Output Set: N:\CRF3\04042001\I509775.raw

```

206 cac tta gcc tgt gat gaa gag aga gtg gaa gag gca aaa ttt ctg gtg 576
207 His Leu Ala Cys Asp Glu Glu Arg Val Glu Glu Ala Lys Phe Leu Val
208      180      185      190
210 act caa gga gca agt att tac att gag aat aaa gaa gaa aag aca ccc 624
211 Thr Gln Gly Ala Ser Ile Tyr Ile Glu Asn Lys Glu Glu Lys Thr Pro
212      195      200      205
214 ctg caa gtt gcc aaa ggg ggc ctg ggt tta ata ctc aag aga cta gca 672
215 Leu Gln Val Ala Lys Gly Gly Leu Gly Leu Ile Leu Lys Arg Leu Ala
216      210      215      220
218 gaa agt gaa gag gct tct atg tag 696
219 Glu Ser Glu Glu Ala Ser Met
220 225      230
223 <210> SEQ ID NO: 4
224 <211> LENGTH: 231
225 <212> TYPE: PRT
226 <213> ORGANISM: Mus sp.
228 <400> SEQUENCE: 4
229 Met Glu Gly Cys Val Ser Asn Ile Met Ile Cys Asn Leu Ala Tyr Ser
230 1      5      10      15
232 Gly Lys Leu Asp Glu Leu Lys Glu Arg Ile Leu Ala Asp Lys Ser Leu
233      20      25      30
235 Ala Thr Arg Thr Asp Gln Asp Ser Arg Thr Ala Leu His Trp Ala Cys
236      35      40      45
238 Ser Ala Gly His Thr Glu Ile Val Glu Phe Leu Leu Gln Leu Gly Val
239      50      55      60
241 Pro Val Asn Asp Lys Asp Asp Ala Gly Trp Ser Pro Leu His Ile Ala
242 65      70      75      80
244 Ala Ser Ala Gly Arg Asp Glu Ile Val Lys Ala Leu Leu Val Lys Gly
245      85      90      95
247 Ala His Val Asn Ser Val Asn Gln Asn Gly Cys Thr Pro Leu His Tyr
248      100      105      110
250 Ala Ala Ser Lys Asn Arg His Glu Ile Ser Val Met Leu Leu Glu Gly
251      115      120      125
253 Gly Ala Asn Pro Asp Ala Lys Asp His Tyr Asp Ala Thr Ala Met His
254      130      135      140
256 Arg Ala Ala Ala Lys Gly Asn Leu Lys Met Val His Ile Leu Leu Phe
257 145      150      155      160
259 Tyr Lys Ala Ser Thr Asn Ile Gln Asp Thr Glu Gly Asn Thr Pro Leu
260      165      170      175
262 His Leu Ala Cys Asp Glu Glu Arg Val Glu Glu Ala Lys Phe Leu Val
263      180      185      190
265 Thr Gln Gly Ala Ser Ile Tyr Ile Glu Asn Lys Glu Glu Lys Thr Pro
266      195      200      205
268 Leu Gln Val Ala Lys Gly Gly Leu Gly Leu Ile Leu Lys Arg Leu Ala
269      210      215      220
271 Glu Ser Glu Glu Ala Ser Met
272 225      230
275 <210> SEQ ID NO: 5
276 <211> LENGTH: 696

```

RAW SEQUENCE LISTING

DATE: 04/05/2001

PATENT APPLICATION: US/09/509,775

TIME: 11:09:33

Input Set : A:\53466277.app

Output Set: N:\CRF3\04042001\I509775.raw

```

277 <212> TYPE: DNA
278 <213> ORGANISM: Rattus sp.
280 <220> FEATURE:
281 <221> NAME/KEY: CDS
282 <222> LOCATION: (1)..(693)
284 <400> SEQUENCE: 5
285 atg gag ggg tgt gtg tct aac cta atg gtc tgt aac ctg gcc tac aac 48
286 Met Glu Gly Cys Val Ser Asn Leu Met Val Cys Asn Leu Ala Tyr Asn
287 1 5 10 15
289 ggg aag ctg gat gag ttg aag gaa agc att ttg gct gat aag tct ctg 96
290 Gly Lys Leu Asp Glu Leu Lys Glu Ser Ile Leu Ala Asp Lys Ser Leu
291 20 25 30
293 gcc act aga act gat cag gac agc aga aca gca ttg cac tgg gca tgc 144
294 Ala Thr Arg Thr Asp Gln Asp Ser Arg Thr Ala Leu His Trp Ala Cys
295 35 40 45
297 tca gct ggt cat aca gaa att gtt gaa ttc ttg ctg caa ctt gga gtg 192
298 Ser Ala Gly His Thr Glu Ile Val Glu Phe Leu Leu Gln Leu Gly Val
299 50 55 60
301 cca gta aat gaa aaa gac gat gca ggt tgg tct cct ctt cat att gct 240
302 Pro Val Asn Glu Lys Asp Asp Ala Gly Trp Ser Pro Leu His Ile Ala
303 65 70 75 80
305 gct tcc gct gcc cgg gat gag att gta aaa gcc ctt ctg ata aaa ggg 288
306 Ala Ser Ala Gly Arg Asp Glu Ile Val Lys Ala Leu Leu Ile Lys Gly
307 85 90 95
309 gca caa gtg aat gcc gtc aat cag aat ggc tgc acg gcc ctt cat tat 336
310 Ala Gln Val Asn Ala Val Asn Gln Asn Gly Cys Thr Ala Leu His Tyr
311 100 105 110
313 gca gct tcc aag aat agg cat gag att gct gtt atg tta cta gaa ggt 384
314 Ala Ala Ser Lys Asn Arg His Glu Ile Ala Val Met Leu Leu Glu Gly
315 115 120 125
317 ggg gct aat cca gat gct aag aac cat tat gat gct aca gca atg cac 432
318 Gly Ala Asn Pro Asp Ala Lys Asn His Tyr Asp Ala Thr Ala Met His
319 130 135 140
321 cgg gca gca gcc aag ggt aac ttg aag atg gtt cat atc ctt ctg ttc 480
322 Arg Ala Ala Ala Lys Gly Asn Leu Lys Met Val His Ile Leu Leu Phe
323 145 150 155 160
325 taa aaa gca tcc aca aac atc caa gat aat gag ggt aac aat cct cta 528
326 Tyr Lys Ala Ser Thr Asn Ile Gln Asp Thr Glu Gly Asn Thr Pro Leu
327 165 170 175
329 cac tta gcc tgt gat gag gag aga gtg gaa gaa gca aaa ttg ctg gtg 576
330 His Leu Ala Cys Asp Glu Glu Arg Val Glu Glu Ala Lys Leu Leu Val
331 180 185 190
333 acc caa gga gca agt att tac att gaa aat aag gaa gaa aag aca ccg 624
334 Thr Gln Gly Ala Ser Ile Tyr Ile Glu Asn Lys Glu Glu Lys Thr Pro
335 195 200 205
337 ctg caa gtc gcc aaa ggg ggc ctg ggt tta ata ctc aaa aga atc gca 672
338 Leu Gln Val Ala Lys Gly Gly Leu Gly Leu Ile Leu Lys Arg Ile Ala
339 210 215 220
341 gaa agt gaa gag gct tct atg tag 696

```

FYI

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY

DATE: 04/05/2001

PATENT APPLICATION: US/09/509,775

TIME: 11:09:34

Input Set : A:\53466277.app

Output Set: N:\CRF3\04042001\I509775.raw

L:178 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:526 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10
L:529 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10